DOCUMENT RESUME

ED 046 898 SP 004 647

AUTHOR Tennyson, Robert D.; And Others
TITLE Exemplar and Nonexemplar Variables Which Produce Correct Concept Classification Behavior and Specified Classification Errors.
PUB DATE 71
NOTE 30p.; Paper presented at annual meeting, AERA, New York, 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Classification, College Students, *Concept Formation, *Concept Teaching

ABSTRACT

Four instructional strategies for promoting the acquisition of infinite concept classes were investigated. Three independent variables were 1) probability level of exemplars and nonexemplars determined by the number of students in a separate sample who correctly classify the instance as an exemplar or a nonexemplar; 2) matching of an exemplar to a nonexemplar so that the irrelevant attributes are the same or very similar; and 3) divergency of an exemplar with another exemplar so that all of their irrelevant attributes differ. Four dependent variables were predicted: 1) correct classification, all instances, exemplars and nonexemplars, correctly identified; 2) overgeneralization, nonexemplars similar to class members identified as exemplars; 3) undergeneralization, low probability exemplars identified as nonexemplars; and 4) misconception, exemplars and nonexemplars sharing a common irrelevant attribute identified as not class members. The four strategies consisted of presenting to S (N=76) a definition followed by 16 exemplars and nonexemplars which were selected according to the hypotheses: 1) IF high to low probability, divergent, and matched THEN correct classification. 2) IF low probability, divergent, and not matching THEN overgeneralization. 3) IF high probability, divergent, and matching THEN undergeneralization. 4) IF high to low probability, convergent, and no matching THEN misconception. A score on each dependent variable was determined for each subject on a specially constructed test requiring S to identify 30 instances of exemplars and nonexemplars. Every hypothesis was supported ($p<.01$). (Author)

EXEMPLAR AND NONEXEMPLAR VARIABLES WHICH PRODUCE CORRECT CONCEPT CLASSIFICATION BEHAVIOR AND SPECIFIED CLASSIFICATION ERRORS

Robert D. Tennyson F. Ross Woolley

M. David Merrill

Brigham Young University

A new concept is acquired when a person correctly identifies previously unencountered objects or events (or representations of such objects or events) as members or nonmembers of a particular class. Concept acquisition can be infered only from situtations where the person is presented unencountered instances and asked to identify those which belong to a given class.

Controversy has resulted in concept research concerning the value of negative instances (nonexemplars) and their relationship to positive instances (exemplars)in promoting concept acquisition. Smoke (1933) concluded that negative instances were of no value in concept learning. Morrisett and Hovland (1959), in replication of Adams' (1954) study of single task vs. multiple task, found that a variety of positive instances was necessary to effect a transfer of concept learning. In studies of combined instances, the equivalent attributes of positive and negative instances are found to be poorly utilized by human subjects, (Bruner, Goodnow, and Austin, 1956; Donaldson, 1959; Hovland and Weiss, 1953). Callentine and Warren (1955) studied positive instances and concluded that the repetition of one or two instances increased attainment. Luborsky (1945) indicated eight exposures was more effective than three.

The research cited deals with concept attainment rather than concept acquisition. In concept attainment studies a finite class is defined. A finite class is one in which all of the exemplars and nonexemplars used in the study are specified. For concept attainment the S could identify all members of the finite class if he were told the relevant attribute (s) (definition). The number of irrelevant attributes associated with a given exemplar of a finite class are limited and carefully controlled. The procedure is inductive in that S is shown instances, either simultaneously or sequentially, and asked to classify each as an exemplar or a nonexemplar. Depending on the variables investigated he may also be able to identify the relevant attributes used to define the class. Usually the same instances, both exemplars and nonexemplars, are presented over and over until S has attained the concept to a given criterion level. This means that it is not possible to determine whether S has memorized mistakes as exemplars or nonexemplars or whether he is transfering his ability to recognize new instances.

In contrast to concept attainment, concept acquisition deals with infinite concept classes. An infinite class is one in which all of the irrelevant attributes associated with a given exemplar cannot be specified. The procedure for presentation is deductive in that S is told what are relevant attributes and given already identified exemplars and nonexemplars prior to the criterion task of identifying class membership of unencountered exemplars and nonexemplars. This process of concept acquisition is defined as a transfer task, (i. e., S will be able to identify class membership of unencountered instances). Once an instance has been presented and identified by the S, it is no longer useful as an item to measure this behavior, hence the need for infinite class.

Mechner (1965) defined concept acquisition as generalization within a class and discrimination between classes. He pointed out that unless both processes were assessed simultaneously it was not possible to infer concept acquisition. In order to assess concept acquisition both exemplars and nonexemplars must be presented to the S and his ability to generalize to new exemplars and discriminate them from nonexemplars is observed. Merrill (1970b) and Markle and Tiemann (1969) postulated that adequate concept acquisition would result only if exemplars were used during instruction which differed widely in the irrelevant attributes associated with each; this promotes generalization within the class. Also, discrimination between classes results from presenting nonexemplars which had irrelevant attributes resembling those associated with given exemplars.

Markle and Tiemann (1969) also postulated that unless the above conditions were met, certain classification behavior errors would result. These are: overgeneralization, undergeneralization, and misconception. Overgeneralization occurs when S correctly identifies all of the exemplars as class members, plus identifying some nonexemplars as members of the class, i. e., the S fails to discriminate between classes. Undergeneralization occurs when S identifies the more obvious exemplars as class members but indicates that less obvious exemplars are not class members, i. e., he fails to generalize to all members of the class. A misconception results when S falsely assumes that some irrelevant attribute or combination of irrelevant attributes is relevant. The operational consequence is that S fails to recognize exemplars not having this attribute as class members and indicates that nonexemplars which do have this attribute are class members.

Woolley and Tennyson (1970) suggested that a more precise operational definition for concept acquisition would result if all exemplars and nonexemplars to be used for a concept class were empirically rated on their probability of being correctly identified by the S when given only a definition (list of relevant attributes). For infinite concept classes the resulting distribution would approximate the normal curve. Their report rated exemplars and nonexemplars on a range from high probability (those exemplars and/or nonexemplars correctly classified by one-half of the Ss as members of a given class) to low probability, (those exemplars and/or nonexemplars which are not correctly classified by one-half of the Ss).

Independent Variables

Based on the theoretical work of Merrill (1970b), Markle and Tiemann (1969, 1970), and Woolley and Tennyson (1970), three independent variables were identified and manipulated in this study:

1) Probability: All exemplars and nonexemplars, preceded with a definition of the relevant attributes, were presented to Ss. High probability items were those instances correctly classified by 60% or more of the sample; medium probability were those correctly classified by more than 30% but less than 60%; and low probability were those instances correctly classified by less than 30% of the sample.

2) Matching: An exemplar and nonexemplar are matched when the irrelevant attributes of the two are as similar as possible. An unmatched relationship between exemplar and nonexemplar occurs when the irrelevant attributes of the two are as different as possible.

3) Divergency: Two exemplars are divergent when the irrelevant attributes of the exemplars are as different as possible. This relationship assumes the same probability level. A convergent relationship occurs when the irrelevant attributes are as similar as possible.

Hypothesis

The three independent variables were combined to predict four dependent variable outcomes. The independent variables refer to characteristics of exemplars presented to S along with a definition. The predicted outcomes were measured using additional unencountered exemplars and nonexemplars which S was asked to classify without confirmation.

Insert Figure 1 about here

The hypotheses are summarized in Fig. 1 and by the following statements:

1) If instances represent a range of probability, exemplars are matched to nonexemplars, and exemplars are divergent with each other, then Ss will correctly classify previously unencountered instances.
2) If instances are low probability, exemplars are not matched to nonexemplars, and exemplars are divergent with each other, then Ss will tend to overgeneralize when classifying previously unencountered instances.
3) If instances are high probability, exemplars are matched to nonexemplars, and exemplars are divergent with each other, then Ss will tend to undergeneralize when classifying previously unencountered instances.
4) If instances represent a range of probability, exemplars are not matched to nonexemplars, and exemplars are convergent with each other, then Ss will tend to demonstrate a misconception when classifying new unencountered instances.

METHOD

Subjects

Thirty-five spring semester junior class educational psychology students enrolled at Brigham Young University served as Ss for the task analysis. Educational psychology classes provided the additional 76 Ss who participated in the study itself. Each S's GPA was used as a covariate.

Program

A poetry concept was selected as the task because it is generally used in literature classroom curriculum, and because the irrelevant attributes of poems are infinite. The concept class taught was trochaic meter. A quality control (Tennyson and Boutwell, 1970) procedure was utilized in developing the four treatment programs: correct classification (to be referred to as classification), undergeneralization, overgeneralization, and misconception. Prototype programs were first organized according to the arrangements of the independent variables (Fig. 1). Subjective knowledge on the part of a subject matter expert determined the various combinations of exemplars and nonexemplars for these first programs.

To determine probability all the program and test selections plus a dozen more instances, 85 total, were randomly scrambled and presented to a group of 35 Ss. The Ss received the definition and the instructions to identify selections they thought to be examples of trochaic meter. From the results the difficulty of the items were determined according to the probability of classification. Fig. 2 shows the percentage of Ss who identified an exemplar to be an example of trochaic meter. The dotted line (Fig. 2) is the percentages of Ss who did not choose a nonexemplar to be an example. Low probability items were those in the 30% level and below. High probability instances were 60% level and above. Items in the middle range were used in the classification and misconception programs. From the above item probabilities the subjective programs and test were revised to fit the appropriate program models. The programs were given to several Ss who then commented individually on the program. Several changes were made as a result, e.g., the instructions were not totally clear, the definition was shortened by eliminating the attribute definition, and several editorial changes.

Insert Figure 2 about here

The programs for the four treatments followed the same format display. Upon concluding the general directions the Ss turned to page 1 and began the self-instructional program. The page 1 directions consisted of a short explanation of classifying examples and nonexamples as a concept of poetry. The following definition was presented:

Part of the rhythm of a poem is determined by the time between stresses occupied with unstressed syllables or pauses. Denoting the stress patterns is to establish the meter. One of the major meter scansions is named trochee and consists of a stressed syllable followed by an unstressed syllable (marked thusly:).

The only instruction for the rest of the program was the presentation of 16 poetry selections. The format of the exemplar/nonexemplar displays consisted of eight pairs of poetry--two each per page. The hypothesized variables were used in constructing the four programs.

In the classification program the exemplars were divergent in their irrelevant attributes, i.e., rhyme, feet, length, style, author, period, etc., arranged from high to low probability, and the nonexemplars were matched to the exemplars in their irrelevant attributes and with similar probability. A sample of the first page of selections for the classification program shows high probability exemplars and nonexemplars:

Divergent	Example:	Matched	Out of childhood into manhood Now had grown my Hiawatha (Longfellow)
	Not an Example:		Come to the crag where the beacon is blazing Come with the buckler, the lance, and the bow, (Scott)
	Example:	Matched	Pansies, lilies, kingcups, daisies. (Wordsworth)
	Not an Example:		Motherly, Fatherly, Sisterly, Brotherly! (unknown)

The classification program selections continually increased in difficulty on each succeeding page.

The overgeneralization task was constructed with divergent low probability exemplars and unmatched with nonexemplars on all four pages. The first page of this program is given here to contrast it with the classification program above.

Divergent

Unmatched

Example:

There they are, my fifty men and women,
Naming me the fifty poems finished!
(R. Browning)

Not an Example:

If the heart of a man is depressed with cares,
The mist is dispell'd when a woman appears.
(Gay)

Unmatched

Example:

Boys in sporadic, tenacious droves
Come with sticks, as certainly as Autumn.
(Eberhart)

Not an Example:

The God of love my Shepherd is,
and He that doth me feed,
While He is Mine, and I am His,
What can I want or need?
(Herbert)

The two exemplars are low probability and divergent; however, the exemplars are unmatched to the nonexemplars. Each page in this program had the same format.

The undergeneralization task had only divergent high probability exemplars which were matched with the nonexemplars. The first page of this program was the same as the classification program, with the succeeding pages on an equal level of difficulty. The example shown here is the last page:

Divergent

Matched

Example: Out of friendship came the Redman
Teaching settlers where the deer ran.
(Imitation, Longfellow)

Not an Example: The smiles that win, the tints that glow,
But tell of days in goodness spent.
(Byron)

Matched

Example: Maid of Athens, 'ere we part,
Give, oh give me back my heart!
(Byron)

Not an Example: Sure solacer of human cares,
And sweeter hope, when hope despairs!
(Bronte)

These exemplars and nonexemplars have a similar probability rating to those illustrated for the high probability classification program.

For the misconception program the convergent grouping was Victorian period trochaic meter poetry. The selections included all three probability ratings of high, medium and low. The nonexemplars were unmatched to the exemplars. Following is an example for this program:

Convergent

Unmatched

Example: There they are, my fifty men and women,
Naming me the fifty poems finished!
(R. Browning)

Not an Example: Give every man thy ear, but few thy voice.
This above all; to thine own self be true.
(Shakespeare)

Unmatched

Example: Wailing, Wailing, Wailing, the wind over
land and sea --
(Tennyson)

Not an Example: When I was one-and-twenty
I heard a wise man say,
"Give crowns and pounds and guineas,
But not your heart away."
(Housman)

The programs were printed and stapled together in a self-instructional booklet. Directions were read aloud by E while the Ss read silently. Once the task began, no questions concerning the program were answered by E. Directions required that Ss not return to previous pages, and since the program was nonspeeded, an

S could spend as much time per page as desired. The test was attached at the end of the program. The width of the test sheets were shortened two inches so that the answers could be placed directly on the answer sheet underneath. The Ss were directed to leave quietly when finished since it was anticipated that Ss receiving the undergeneralization task would finish before others.

Test

The test was constructed so that the predicted responses of the dependent variables could be analyzed. Thirty selections of poetry were sectioned into three parts with the following format:

1. Convergent high probability exemplar.
2. Convergent low probability exemplar.
3. High probability nonexemplar matched to number 1.
4. Low probability nonexemplar matched to number 2.
5. High probability nonexemplar unmatched.
6. Divergent high probability exemplar paired to number 1.
7. Divergent low probability exemplar paired to number 2.
8. High probability nonexemplar matched to number 6.
9. Low probability nonexemplar matched to number 7.
10. Low probability nonexemplar unmatched.

The thirty selections were randomly scrambled so that no patterns were evident to the Ss. To test the dependent variable of misconception the grouping of Victorian period poetry was identified as convergent, all other grouping was classified as divergent. Ss in the misconception treatment condition were hypothesized to classify only convergent high and low probability exemplars and identify as exemplars those matched nonexemplars as exemplars. The classification treatment group was hypothesized to correctly classify all exemplars on the test. The overgeneralization treatment group was hypothesized to classify not only the exemplars but the low probability matched nonexemplars as instances. They could also pick high probability matched and unmatched nonexemplars and not be penalized by an error, i. e., an S in this group could have classified all 30 items as exemplars and still follow the predicted results. Undergeneralization Ss were hypothesized to respond only to the high probability exemplars.

The hypothesized response patterns for each of the dependent variables are given in Fig. 3. S responses were compared with the predicted score for each dependent variable. S was scored with an error for a given dependent variable when his response to a given item differed from the predicted response. Scores were obtained for the three sections of the test and then added together for the four separate dependent variable conditions. This procedure gave the S four scores; one for each hypothesized dependent variable.

Insert Figure 3 about here

Experimental Design

A posttest-only control group experimental design was used in this study

(Campbell and Stanley, 1966). Inter ·l validity was controlled by random assignment of Ss to the four programs. Since the programs were administered to individual Ss the basic experimental unit was the S. The n-size for each treatment was 19, total n=76. External validity is a problem since the Ss were not randomly sampled from the universal population. With replication of the study external validity will be strengthened.

RESULTS

Variable Measures

An analysis of covariance was used for several reasons. The first was to increase precision on a randomized experiment. The covariate of GPA was a measure of each experimental unit before the treatments were applied. The assumptions for analysis of variance were made, along with the assumptions of covariance. Test for homogeniety of regression of within-class and between-class linearity was performed meeting this assumption. The second reason was to throw light on the nature of the treatment effects. There are numerous instances in which a concomitant variable might be in part the agent through which the treatments produce their effects on the principal response. A covariance analysis offers the possibility of exploring whether this is so. Thirdly, an analysis of covariance was used to fit regressions in multiple classifications. By standard techniques we could (i) fit a separate regression of Y on X within each group, (ii) test whether the slopes or positions of the lines differ from group to group, and (iii) if desired, make a combined estimate of a common slope. Only the adjusted means from the analysis of covariance are presented since the unadjusted means did vary significantly from the adjusted means.

Four error scores were obtained for each S according to the predicted responses on the dependent variables (Fig. 3). S's responses were compared to the predicted dependent variables and were scored one error for each deviation. The four scores were based not on responses to correct answers but on predicted responses as the results of the manipulation of the three independent variables in the four tasks. Fig. 4 shows the treatment groups, represented by capital letters, and the predicted errors for each dependent variable, i.e., the C group would make zero errors under the correct classification variable but it was predicted that O (overgeneralization) group would make eight errors, the U (undergeneralization) group six errors, and the M (misconception) group would make nine errors. Thus each group was predicted to make significantly fewer errors than the other three conditions when its dependent variable was analyzed. Likewise, the other variations in error scores per group were predicted.

Insert Figure 4 about here

The adjusted covariate means for the four treatment groups according to the dependent variables are listed in Table 1. A separate analysis of covariance was used for each dependent variable, i.e., for the classification variable the means used were C (5.68), O (12.97), U (9.83), and M(11.98). A multivariate analysis of covariance was not appropriate because the S had only one measured response per item. For each dependent variable the S was measured according to the predicted criterion of that variable. There was a difference between means for all four treatments ($p < .01$). The four covariate F tests (3, 72 df) were: Classification on F=16.65; Overgeneralization, F=12.57; Undergeneralization, =24.79; and Misconception, F=7.44. The posteriori tests used to determine which individual means were significantly different from one another were the Newman-Keuls Sequential Test and Duncan's New Multiple Range Test.

Insert Table 1 about here

Correct Classification

The hypothesized correct classification variable was constructed of matched exemplars and nonexemplars, and divergent exemplars on a high to low probability continium (Fig. 1). On both the Newman-Keuls and Duncan's the C group made fewer errors than the O, M, and U groups ($p < .01$). This corresponds to the hypothesis and the predicted results in Fig. 4. There was a difference between the O group and the U group on the Newman-Keuls ($p < .05$) and on the Duncan's ($p < .01$). According to Fig. 4 there is a predicted difference of two errors between the U group and the O group. No difference was found between U and M groups ($p > .05$). A three point difference was predicted.

Overgeneralization

The overgeneralization dependent variable resulted from divergent low probability exemplars that were unmatched with nonexemplars. The multiple comparison of the Newman-Keuls showed a difference between the O group and the U group ($p < .01$); this follows the prediction from Fig. 4 of an error spread of 14 points. A difference existed between O group and M and C groups on the Neuman-Keuls ($p < .05$). Duncan shows a difference for O with U and M groups ($p < .01$), and O group and C group ($p < .05$). Other predicted differences from Fig. 4 on both tests were between the C group and U group and the M and U groups ($p < .01$). There was no difference between C group and M group ($p > .05$).

Undergeneralization

The undergeneralization treatment condition received the independent variables of high probability divergent exemplars and matched nonexemplars. The multiple comparisons of the undergeneralization error scores show for both the Newman-Keuls and Duncan's that the U group differs from the O and M groups ($p < .01$). The predicted errors be*ween the U and C groups was six (Fig. 4). There was a .01 level of significance between C and U on the correct classification analysis, but here, the difference was only at the .05 level on both tests. This is probably the result of the U group generalizing more than predicted. Other predicted differences are: the O group received a higher error mean than the other groups on this dependent variable, as predicted ($p < .01$); the difference between O group and M group were predictably lower ($p < .01$); the difference narrows on the C group and the M group ($p < .01$).

Misconception

In the misconception treatment group the Ss were instructed with convergent Victorian period, high, medium, and low probability exemplars with unmatched nonexemplars. The results follow the predicted variables on all factors on both tests ($p < .01$). The M group was different from O, U, and G groups ($p < .01$). No significance resulted from the comparison on the other three groups as predicted in Fig. 4 ($p > .05$).

Discussion

The significant results of this study have definite implications for instructional procedures on the cognitive level of behavior. Instruction is the manipulation of the environment so that student behavior can be changed. If instruction does not include empirically founded principles, the student may not learn all that is desired by the instructor and student. The problems discussed by Markle and Tiemann (1969, 1970) on overgeneralization, undergeneralization, and misconception as a result of faulty classification behavior instruction of a concept class is now more than a hypothetical position.

Precise independent variables were arranged in such a way that predicted dependent variables did result in all cases. Four of the twelve difference predictions (Fig. 4) did not reach the .01 level, but were significant at the .05 level. With some adjustment in the items used, those four predictions could be increased to the .01 level. The implications are clear that instruction does produce certain types of dependent variables and if these are not controlled by empirically based procedures there is little assurance that students learn behavioral objectives--no matter how precisely and Magerian (1962) we may state them.

The results obtained in this study were based on the task analysis which provided the probability ratings. Subjective subject matter expertise is not sufficient to produce the dependent variables reached here. The independent variable of probability rating of exemplars is crucial in instructional design. The task analysis did produce ratings that the author could not have known. Task analysis is not a new procedure (Gagné, 1965, 1968). However, the implications in the context of this research have not been implemented on an empirically based instructional system. The most significant difference obtained in this study was between the undergeneralization group and the overgeneralization group. The undergeneralization group was presented only high probability exemplars and, as a result, responded to few items on the test. On the contrary, the overgeneralization group received only low probability exemplars and responded to practically everything on the test. The significant results were not only for the main effects, but for the predicted differences for the other effects on each dependent variable.

The independent variable of matched exemplars/nonexemplars in an infinite concept is, also, definitely involved. This can be seen empirically by the increased response to nonexemplars (negative instances) by the overgeneralization and misconception groups. In both cases the nonexemplars were unmatched to the exemplars so that Ss failed to recognize the relevant attributes from the irrelevant attributes. Again the principle of probability is involved since the generalization and undergeneralization groups received matching on equal probabilities. In both cases fewer nonexemplars were chosen as exemplars. The implication is that discrimination is more effectively taught if the matching of exemplars and nonexemplars is empirically controlled by a task analysis probability rating of both exemplars and nonexemplars.

The third independent variable of the relationship between exemplars according to their irrelevant attribute groupings was significant. The three treatments of classification, undergeneralization, and overgeneralization all received divergent exemplars based on their probability ratings. Only the misconception group did not receive divergent exemplars. The importance of divergency of equal probability exemplars that differ in all irrelevant attributes as opposed to presenting convergent exemplars is empirically shown by the predicted results of the misconception dependent variable. This treatment group did not choose exemplars which differed from those presented in the misconception task.

More work is needed on different subject matter tasks and sample populations to add external validity to the results gained here. Further research will expand the variables and implications obtained in this study. Areas that need investigation are: the most effective number of exemplars and nonexemplars; modification procedures to correct the three problems of concept instruction; the effect of the three problems on higher levels of cognitive learning; and paralleling instructional strategies and independent variables on the higher levels of learning.

REFERENCES

Adams, J. A. Multiple versus single problem training in human problem solving. J. exp. Psychol., 1954, 48, 15-18.

Bruner, J. S., Goodnow, Jacqueling, J., & Austin, G. A. A study of thinking. New York: Wiley, 1956.

Callentine, M. F., & Warren, J. M. Learning sets in human concept formation. Psychol. Rep., 1955, 1, 363-367.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1968.

Donaldson, M. Positive and negative information in matching problems. Brit. J. Psychol., 1959, 50, 253-262.

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart & Wilson, 1965.

Gagné, R. M. Learning hierarchies. Paper presented at the meeting of the American Psychological Association, San Francisco, August, 1968.

Hovland, C. I., & Weiss, W. Transmission of information concerning concepts through positive and negative instances. J. exp. Psychol., 1953, 45, 175-182.

Luborsky, L. F. Aircraft recognition: I. the relative efficiency of teaching procedures. J. appl. Psychol., 1945, 29, 385-398.

Mager, R. F. Preparing instructional objectives. Palo Alto, Calif.: Fearon Pub., 1962.

Markle, S. M., & Tiemann, P. W. Really understanding concepts. Champaign, Ill.: Stipes, 1969.

Mechner, F. Science education and behavioral technology. In R. Glaser (Ed.), Teaching machines and programed learning, II: Data and directions. Washington D. C.: National Education Association, 1966, 441-507.

Merrill, M. D. Understanding the instructional objective to understand, Tape-slide lecture presented MSPI mtgs, Washington D. C., 1969. MSPI mtgs, Disneyland, Calif., 1970.

Merrill, M. D. (Ed.) Readings in instructional design. Englewood Cliffs, N. J.: Prentice-Hall, 1970a, in press.

Merrill, M. D. Necessary psychological conditions for defining instructional outcomes. Educational Technology, 1970b, in press.

Morrisett, L. Jr. & Hovland, C. I. A comparison of three varieties of training in human problem solving. J. exp. Psychol., 1959, 58, 52-55.

Smoke, K. L. An objective study of concept formation. Psychol. Monogr., 1932, 42 (4, Whole No. 191).

Tennyson, R. D. & Boutwell, R. C. A quality control design and evaluation model for hierarchical sequencing of programed instruction. NSPI Journal, 1970, in press.

Woolley, F. R. & Tennyson, R. D. Conceptual model of classification behavior. Educational Technology, 1970, in press.

FOOTNOTE

[1]This project was supported by the United States Office of Education Small Contract No. O-H-014 and Brigham Young University Department of Instructional Research and Development. The authors would like to express their appreciation to Mrs. Carol L. Tennyson for her part as the subject matter expert in preparing the task used in this study. Thanks also goes to our research assistant, Miss Lynda J. Crosby who helped in the analysis. A report of this research was presented at the American Educational Research Association 1971 annual meeting.

TABLE 1

Mean Scores of the Four Dependent Variables

		Means			
	Groups	C	O	U	M
Dependent Variables	Class.	5.68	12.97	9.83	11.98
	Over.	9.01	7.00	11.83	9.25
	Und.	8.55	14.33	6.22	11.62
	Mis.	9.80	10.52	9.75	7.38

Dependent Variable Outcome		Independent Variables Presented		
		Probability	Matching	Divergency
	Correct Classification	All Levels	Matched	Divergent
	Over-generalization	Low Or All Levels	Unmatched	Divergent
	Under-generalization	High Level	Matched Or Unmatched	Divergent
	Misconception	All Levels	Unmatched	Convergent

Fig. 1 The hypotheses matrix.

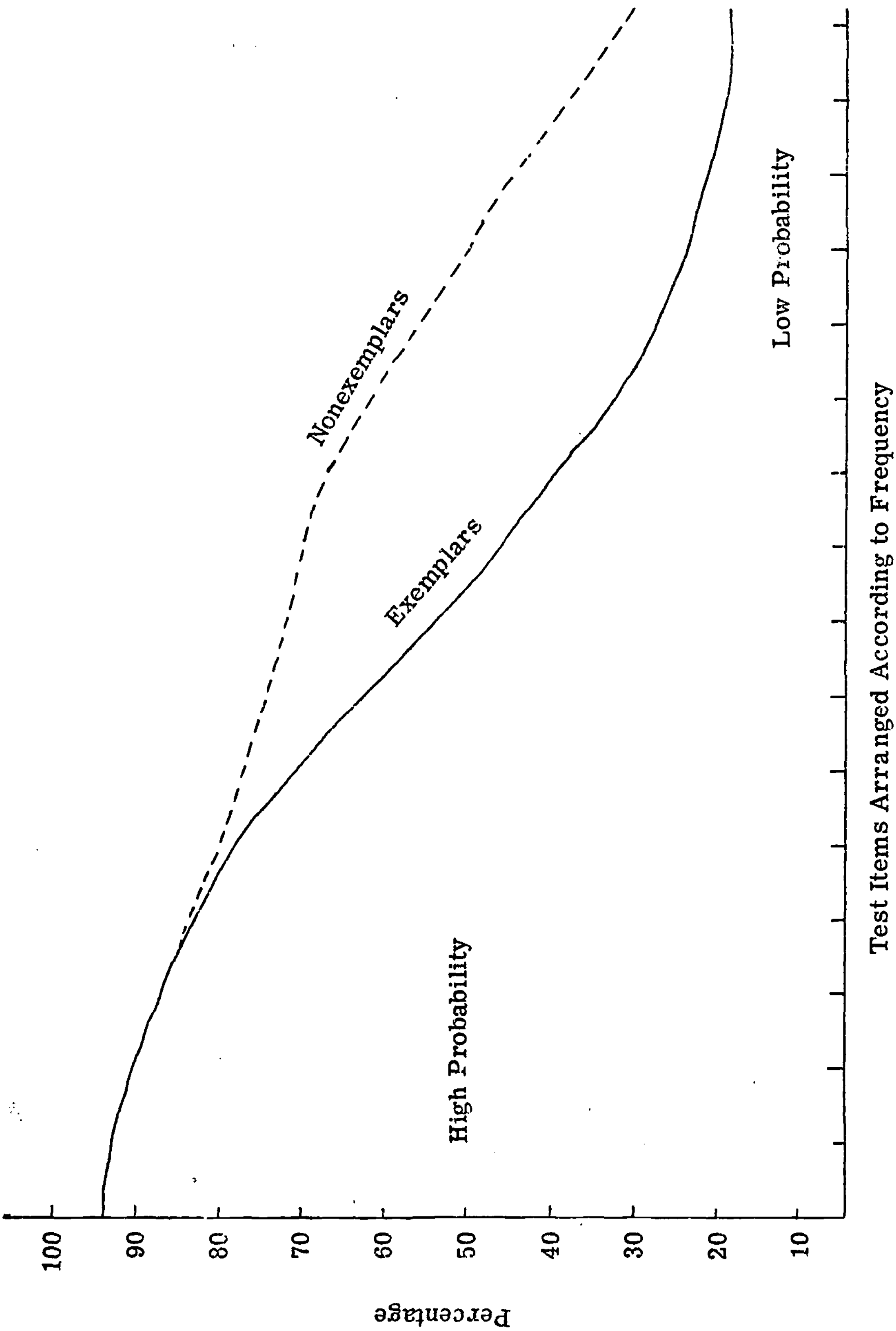

Fig. 2. Frequency distribution of exemplars and nonexemplars used on the task analysis to determine probability.

		M	O	U	G
1.	eg #6	X	X	X	X
2.	eg #16	X	X	O	X
3.	$\overline{eg}$ #3	-	-	O	O
4.	$\overline{eg}$ #30	X	X	O	O
5.	$\overline{eg}$. #15	-	-	O	O

		M	O	U	G
6.	eg #12	O	X	X	X
7.	eg #21	O	X	O	X
8.	$\overline{eg}$ #8	O	-	O	O
9.	$\overline{eg}$ #17	O	X	O	O
10.	$\overline{eg}$ #4	O	X	O	O

Fig. 3. Scoring sheet. Predicted responses according to conditions.

Note. --M = misconception; O = overgeneralization; U = undergeneralization; C = correct classification; X = S indicates this selection is an exemplar; O = S indicates this selection is a nonexemplar; - = S could classify as either, no error possible; eg indicates an exemplar; $\overline{eg}$ indicates a nonexemplar; # refers to original test item number.

Predicted Errors

Dependent Variables	Groups	C	O	U	M
	Class.	0	8	6	9
	Over.	3	0	14	11
	Und.	6	14	0	9
	Mis.	9	11	9	0

Fig. 4. Hypothesized error responses.